Fuzzy Clustering of Homogeneous Decision Making Units with Common Weights in Data Envelopment Analysis

Sajad Kazemi*a, Reza Kiani Mavi*b, Ali Emrouznejadc, Neda Kiani Mavi*b

a Graduate School of Management, Saint Petersburg State University, St. Petersburg 199004, Russia.
b School of Business and Law, Edith Cowan University, Western Australia, Australia.
c Aston Business School, Aston University, Birmingham, UK.

Abstract. Data Envelopment Analysis (DEA) is the most popular mathematical approach to assess efficiency of decision-making units (DMUs). In complex organizations, DMUs face a heterogeneous condition regarding environmental factors which affect their efficiencies. When there are a large number of objects, non-homogeneity of DMUs significantly influences their efficiency scores that leads to unfair ranking of DMUs. The aim of this study is to deal with non-homogeneous DMUs by implementing a clustering technique for further efficiency analysis. This paper proposes a common set of weights (CSW) model with ideal point method to develop an identical weight vector for all DMUs. This study proposes a framework to measuring efficiency of complex organizations, such as banks, that have several operational styles or various objectives. The proposed framework helps managers and decision makers (1) to identify environmental components influencing the efficiency of DMUs, (2) to use a fuzzy equivalence relation approach proposed here to cluster the DMUs to homogenized groups, (3) to produce a common set of weights (CSWs) for all DMUs with the model developed here that considers fuzzy data within each cluster, and finally (4) to calculate the efficiency score and overall ranking of DMUs within each cluster.

Keywords: Data Envelopment Analysis, Fuzzy DEA, Non-homogeneous, Clustering, Common Set of Weights (CSW).

1. Introduction

Banking system is the pillar of the economy in most countries and has a significant role in developing the economic and financial systems by mobilizing small deposits of people in the household sector and directing them towards productive uses in the industrial sectors. For example, growth and development of the banking industry and financial markets pave the way for many organizations to be more innovative by receiving sufficient finance from the banking sector and improving the operational efficiency [1]. Therefore, evaluating the performance of the banking system helps to achieve higher economic growth [2]. There is wide ranging research available that has looked at the efficiency of the commercial banking industry [3, 4]. Banks need to boost their efficiency to remain competitive. Analyzing efficiency stimulates the banks to move forward by looking at the best performing branches [5]. A citation-based systematic literature review on banking sector performance by Ahmad et al. [6] demonstrates that studies in this realm widely have employed stochastic frontier analysis (SFA) as a parametric approach, whereas data envelopment analysis (DEA) is prevalent in nonparametric approach to measure the banking sector performance and efficiency. For example, recent studies such as Tavana et al. [7], Barros and Wanke [5] and Kiani Mavi et al. [8] have focused on the efficiency analysis using DEA models.

There are basic assumptions in DEA such as measuring inputs and outputs with real values and definite
factors, while in practice the values of inputs and outputs might include missing data, judgmental data, forecasted data, or generally imprecise data [9]. Assigning a crisp value for a judgmental variable is difficult because of ambiguity, limitations, and complexity of them. To deal with this shortcoming, fuzzy DEA models have been developed in which data of inputs/outputs are expressed by fuzzy linguistic variables [10]. Hatami-Marbini et al. [11] presented a taxonomy and review of the fuzzy DEA methods and presented a classification scheme in this regard.

On the other hand, conventional models of DEA split decision-making units (DMUs) into two categories of efficient DMUs with the efficiency score of 1 and inefficient DMUs with the efficiency score of a value between [0,1). Thus, conventional models do not distinguish efficient units and cannot rank them. Several methods including common set of weights (CSW) have been developed to completely rank the efficient DMUs in DEA. The flexibility in determining the weights by each DMU to maximize its efficiency score makes it difficult or even sometimes impossible to rank all DMUs based on the efficiency scores especially when many DMUs obtain the efficiency score of one or very close to one which leads to low discrimination among DMUs [12]. Thus, CSW procedure diminishes the disbandment of the optimal weights allocated to the input and output variables by each DMU. The CSW approach essentially obtains a common compromised set of weights in lieu of different sets of weights for assessing the efficiency of DMUs, so evaluates the efficiency of DMUs with a common set of weights. Therefore, DMUs are not allowed to allocate different weights to their inputs and outputs. Common weight models make it possible to compare and rank all DMUs with the same optimal weights [13; 14].

Conventional DEA models accept that all DMUs are homogenous and operate in a similar operating environment [8]. However, in real-world situations, different environmental variables challenge homogeneity of DMUs, DEA studies have paid relatively less attention to them. Athanassopoulos [15] points out that while bank branches are similar in the operations, they meet different environments in the market. Therefore, not considering the factors outside the control of branch manager will reduce their efficiency score in the performance evaluation plan. So, factors that make branches different should be taken into account and similar branches should be clustered to achieve homogeneous sets. Then, performance evaluation of DMUs should take place to the homogeneous set of DMUs. In addition, to deal with the high number of decision makers and performance criteria which cause higher complexity and vagueness of decision-making process, clustering of DMUs is essential [16]. Overall, clustering and DEA are integrated to each other via two different approaches. First, the results of clustering are integrated to the classification resulted from DEA to create several reference subsets from the initial reference set of DMUs. Second, the efficiency score of a DMU is not defined by the peer group, but it is defined by an efficient subset of the same peer subset. Subsequently, this method reaches multiple homogeneous subsets by measuring similarity among units, thus, every DMU is compared with its peers in the given subset [17].

Considering the above-mentioned issues, this study contributes to the existent literature by providing a new mathematical model to resolve them. Therefore, major contributions of this paper are as follows:

- This paper proposes a novel common weight model based on ideal point method (CSW-IP) under the fuzzy environment to overcome the problems of ranking efficient DMUs and fuzzy data altogether.
- The proposed approach can easily define the best performing DMU while traditional DEA models are not able to do so.
- This study identifies the environmental factors that influence the efficiency of bank branches. Then, the efficiency of bank branches in each cluster is measured utilizing the proposed model to provide more realistic results.

The remainder of this paper is organized as follows. Section 2 reviews the literature of DEA and clustering method. Section 3 presents the proposed CSW-F-IP model. Section 4 illustrates a real case to show the applicability of the proposed model. Section 5 discusses the advantages of the proposed approach. Finally, conclusion and direction for future studies are drawn in Section 6.

2. Literature review

2.1. Non-Homogeneity among DMUs in Efficiency Analysis

Data envelopment analysis (DEA) is a technique to evaluate the relative efficiency of a set of homogeneous decision-making units (DMUs). Homogeneity means that all DMUs utilize the same inputs to generate the same outputs in different amounts from one DMU to another [18]. For example, manufacturing plants in the same industry as DMUs might not all
utilize similar inputs to produce the same products. Fair comparison of DMUs in terms of efficiency analysis in the absence of homogeneity is an issue for real world situations [19]. By homogeneity, all DMUs must obey three conditions: first, DMUs should execute the same processes; second, their efficiency should be evaluated by the same input and output variables means that DEA model applies equally to all DMUs; third, all DMUs operate within the same environment under the same conditions [20]. However, in real world practice, these hypothetical assumptions are not always applied. Undoubtedly, to compare performance of two bank branches in different cities, decision makers will face environmental heterogeneity that should be considered as an environmental variable, which reveals economic, societal, political, environmental, legal, and cultural differences among different environments. For example, Dyson et al. [21] and Athanassopoulos [15] have emphasized on the environmental variables especially at service sectors in which identification, definition, and evaluation of these variables are far more difficult. Heterogenous preferences in group decision making problems may also stem from decision makers. Chen et al. [22] reviewed three types of fusion approaches addressed the integrating and aggregating the heterogenous preference structures. Despite successfully identification of heterogeneous variables in terms of environmental differences among DMUs, accuracy of the inputs and outputs and efficiency scores is not guaranteed. This issue is more important when decision makers should express their opinion on variables, which adds a new heterogeneity to the model [22].

A variety of DEA-based techniques have been developed to handle different environmental conditions. Haas and Murphy [23] mentioned two approaches to manage non-homogeneity among DMUs. The first approach is to partition DMUs into homogeneous clusters. The second approach is to accommodate for non-homogeneity. Saen et al. [24] used analytic hierarchy process (AHP) chance-constrained DEA for calculating the relative efficiency of slightly non-homogeneous DMUs. Cook et al. [25] proposed hierarchical DEA to handle the non-homogeneity of DMUs by measuring efficiency at multiple levels. They delivered a technique for modifying the ranking of DMUs at one level to account for the rankings obtained by the groups (which these DMUs belong to) at a higher level. It should be noted that they classified non-homogeneous DMUs into different groups and evaluated them in different models. A two-step process is commonly used to manage the non-homogeneity issue in DEA [26; 23]. DEA efficiency values are obtained in the first step, then in the second step, other sources of non-homogeneity that were not considered in the original modeling are taken into account to regress the DEA efficiency scores. When environmental factors cause non-homogeneity, they are considered in a single model as non-discretionary inputs. Therefore, different reference sets are defined to discriminate DMUs in different environments [28]. In DEA, the approaches used to deal with non-homogeneity, are classified as the (i) frontier separation, (ii) all in-one, (iii) two-stage, (iv) environmental harshness index approaches [29]. Although DEA models with non-discretionary factors such as distinguishing the internal and externa non-discretionary variables [30] have been developed, there is no generally accepted approach for reflecting non-discretionary variables in DEA models. Therefore, this study classifies DMUs based on non-discretionary variables then evaluate them with DEA.

2.2. Clustering

Clustering is a subset of multivariate statistical techniques that finds hidden patterns in a set of data. It divides a set of data (operating units) into different subsets during the process of grouping them according to pre-defined criteria so that similar objects are grouped in a distinct cluster. Clustering is usually done by maximizing the similarity of objects within each cluster [31; 32].

Generally, clustering methods are divided into the following groups: hierarchical clustering, mixture-model clustering, learning network clustering, objective-function-based clustering, partition clustering, distance-based clustering, spectral clustering, kernel-based clustering, and graph-based clustering [17; 32]. Distance-based algorithms are the most common algorithms that minimize the overall dissimilarity of objects. They make an initial cluster and then use a repeating movement technique to maximize the overall similarity of objects within clusters or minimize the overall dissimilarity by transferring the objects from one cluster to the other cluster. The most common methods used for this type of clustering includes K-means, fuzzy C-means (FCM), Possibilistic C-means (PCM) and fuzzy equivalence relation (FEC) [17; 32] which C-means algorithm is the most practical approach. In this algorithm, the number of clusters should be pre-identified and the results of classification significantly rely on the selection of the threshold values. The issue for decision makers is the difficulty of judgment about the classification when they have limited information. Identifying the number of
intended clusters in FCM method is difficult. To decrease the complexity of calculations, fuzzy clustering method is more appropriate when uses equivalence relation for dividing DMUs of a large set into multiple smaller sets. Due the existence of ambiguity and limitations of data or complexity of decision-making factors, doing relative preferences with crisp numerical values might lead to inaccurate results. In contrast, modelling linguistic variables with fuzzy sets can flexibly express the relative preferences of decision makers. Therefore, clustering with fuzzy equivalence relation (FEC) has been considered as a method which provides more appropriate classification results, and also solves the problem of clusters’ overlapping by applying max-min transitive closure approach [28]. Samoilenko and Osei-Bryson [23] and Paradi et al. [34] integrated DEA with clustering including.

2.2.1. Fuzzy Equivalence Relation (FEC) Algorithm

Fuzzy equivalence relation method supposes that a system comprises a group of criteria \( C = (C_1, C_2, \ldots, C_m) \), and especial mutual relations are identified to modelling with mathematical relations. We consider a group of experts (executives or managers) who have a substantial role in establishing strategy map for an organization and are knowledgeable about criteria, alternatives, and their influence on the goal, in order to evaluate alternatives against criteria. After gathering experts’ evaluations, all matrices are aggregated as an average matrix. Therefore, the solution procedure of Wang [33] is extended with this point of view:

Step 1. Determine decision objective. Identify the relevant criteria to evaluate the pros and cons of alternatives to ensure the goal is achieved. A group of experts, \( K \), should be formed to evaluate alternatives against criteria to achieve the goal.

Step 2. Generate the initial direct influence matrix. Measuring the influence of each alternative against the criteria requires designing a comparison scale, for example, seven levels of influence as 1 (very low), 2 (low), 3 (between low and medium), 4 (medium), 5 (between medium and high), 6 (high), and 7 (very high) influence. An initial direct influence matrix \( A^k \) is a matrix constructed by the pairwise comparisons done by the expert \( k \), in which \( a_{ji} \) is the extent to which the alternative \( j \) is evaluated against the criterion (variable) \( i \) in order to achieve the goal, i.e. \( A^k = [a_{ji}]_{n \times m} \).

Step 3. Convert the linguistic terms into fuzzy linguistic scale and transform them into fuzzy numbers as shown in Table 1.

<table>
<thead>
<tr>
<th>Influence level</th>
<th>Linguistic terms</th>
<th>Symbol</th>
<th>Fuzzy numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Very low</td>
<td>VL</td>
<td>(0, 0.2)</td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
<td>L</td>
<td>(0, 0.2, 0.4)</td>
</tr>
<tr>
<td>3</td>
<td>Between low and medium</td>
<td>BLM</td>
<td>(0.2, 0.4, 0.5)</td>
</tr>
<tr>
<td>4</td>
<td>Medium</td>
<td>M</td>
<td>(0.4, 0.5, 0.6)</td>
</tr>
<tr>
<td>5</td>
<td>Between medium and high</td>
<td>BMH</td>
<td>(0.5, 0.6, 0.8)</td>
</tr>
<tr>
<td>6</td>
<td>High</td>
<td>H</td>
<td>(0.6, 0.8, 1)</td>
</tr>
<tr>
<td>7</td>
<td>Very high</td>
<td>VH</td>
<td>(0.8, 1, 1)</td>
</tr>
</tbody>
</table>

Step 4. Completing step 3 leads to \( K \) fuzzy initial direct influence matrices shown by \( A^k \), \( k = \{1, 2, \ldots, K\} \) in which \( A^k_{ij} = (a^k_{ij}, a^k_{mj}, b^k_{ij}) \) is the fuzzy preferences of the alternatives against criteria specified by the decision maker \( k \). The average of preferences is calculated by (1) [35] as:

\[
\bar{p} = \frac{\sum_{i=1}^{K} A^k_{ij}}{K}
\]  

Here, a preliminary of mathematics of fuzzy numbers is presented [33]:

Definition 1. A triangular fuzzy number (TFN) is described as a triplet \((a, a_m, a_r)\). If \( \tilde{A} \) is a TFN, its membership function is defined as (2).

\[
f_{\tilde{A}}(X) = \begin{cases} \frac{x-a}{a_m-a_r}, & a_r \leq x \leq a_m, \\ \frac{a_m-x}{a_m-a_r}, & a_m < x \leq a_r, \\ 0, & \text{otherwise} \end{cases}
\]  

Assume \( \tilde{A} \) and \( \tilde{B} \) are two TFNs shown by the triplet \((a_b, a_m, a_r)\) and \((b_l, b_m, b_r)\) respectively. Fuzzy arithmetic operations on \( \tilde{A} \) and \( \tilde{B} \) is as (3-6):

\[
\tilde{A} + \tilde{B} = (a_1 + a_b, a_m, a_r + b_r) 
\]

\[
\tilde{A} - \tilde{B} = (a_1, a_m, a_r - b_r) - (b_l, b_m, b_r) = (a_1 - b_r, a_m - b_m, a_r - b_l) 
\]

\[
\tilde{A} \times \tilde{B} = (a_1, a_m, a_r) \times (b_l, b_m, b_r) = (a_1 \times b_l, a_m \times b_m, a_r \times b_r) 
\]

\[
\tilde{A} = (a_1, a_m, a_r) 
\]

\[
\tilde{B} = (b_l, b_m, b_r) 
\]

where \( a_1, a_m, \) and \( a_r \) are real values and \( a_1 < a_m < a_r \).

Step 5. Let \( \tilde{X}_l = (X_{l1}, X_{l2}, \ldots, X_{ln}) \) denotes ith \((i = 1, 2, \ldots, m)\) sequence comprising of \( n \) TFNs. Assume \( D \) indicates a data matrix constructed of \( m \) fuzzy sequences; so that,

\[
\tilde{D} = [\tilde{X}_{lq}]_{m \times n}
\]
in which $i$ shows the sequence index and $q$ points at the fuzzy number index.

Step 6. Considering the data matrix $D$, the fuzzy binary relation between any two fuzzy sequences is defined as follows:

Assume $X_iq = (\lambda_iq_0, \ldots, \lambda_iq_m, \lambda_iq_n)$ is a triangular fuzzy number $i = 1, 2, \ldots, m$; $q = 1, 2, \ldots, n$. Upper and lower bounds of $q$th TFN of all sequences in matrix $D$ are defined by $\bar{T}_q^a = (t^+_a q, t^-_a q, t^+_a q, t^-_a q)$ and $\bar{T}_q^a = (t^-_a q, t^-_a q, t^+_a q, t^-_a q)$, respectively, where

$$
t^+_a q = \max_i \{ x_{iq} \}, t^-_a q = \max_i \{ x_{iq} \}, t^+_a q = \max_i \{ x_{iq} \}
$$

$$
t^-_a q = \min_i \{ x_{iq} \}, t^-_a q = \min_i \{ x_{iq} \}
$$

$i = 1, 2, \ldots, m$; $q = 1, 2, \ldots, n$.

Step 7. Binary relation between fuzzy sequences $X_i$ and $X_j$ is described by $R = \{(i, j), \mu_R(i, j)\}$. Therefore, the membership $\mu_R(i, j) = 1$ meets reflexive and symmetric conditions, i.e. $\mu_R(i, j) = \mu_R(j, i)$, respectively, therefore, the fuzzy binary relation $R$ will be fuzzy compatible relation.

Definition 2. In the $n \times n$ fuzzy relation matrix $R = [r_a(i, j)]_{n \times n}$, $0 \leq r_a(i, j) \leq 1$ represents the relationship between two sequences $i$ and $j$, and $0 \leq i, j \leq n$. Matrix $A$ is reflexive, if $r_a(i, i) = 1$ for $0 \leq i \leq n$, and it is symmetric, if $r_a(i, j) = r_a(j, i)$. The matrix $A$ is named max-min transitive, if $r_a(i, j) \geq \max_k \min (r_a(i, k), r_a(k, j))$.

Definition 3. When a fuzzy relation matrix satisfies both reflexive and symmetric conditions, then it will be a fuzzy compatible matrix.

Definition 4. A fuzzy equivalence matrix is a fuzzy compatible matrix which meets max-min transitive.

Step 8. $R$ will be reflexive and symmetric but might not be transitive. Therefore, $R$ is just a fuzzy compatible relation, not a fuzzy equivalence relation.

When fuzzy compatible relation is used to clustering, fuzzy data sequences might have some overlaps. To remove the overlap, the compatible relation must be converted into a fuzzy equivalence relation by an approach such as max–min transitive closure. The max–min transitive closure $R^*$ of $R$ is presented as (12–13):

$$
\mu_R(i, j) = \max_k \min [\mu_R(i, k), \mu_R(k, j)]
$$

and

$$
\mu_R(i, j) = \max_k \min [\mu_R(i, k), \mu_R(k, j)]
$$

Integrating above-mentioned relations, leads to the revised $R^*$ and $R^+$:

$$
\mu_R(i, j) = \mu_{R^+}(i, j) \geq \mu_{R^*}(i, j)
$$

where $R^+$ is derived by following definitions if $\mu_{R^*}(i, j) \geq \mu_{R^+}(i, j)$, $\forall i, j$.

Definition 5. An $n \times n$ fuzzy relation matrix $B = [r_B(i, j)]_{n \times n}$ contains the matrix $A$, indicated as $A \subseteq B$, where $r_B(i, j) \leq r_B(i, j)$ for all $i, j$.

Definition 6. When a fuzzy equivalence matrix $B$ contains a fuzzy compatible matrix $A$ and is included by any fuzzy equivalence matrix that contains $A$, then it is called the equivalence closure.

Step 9. In addition, $R^*$, based on $R$ and meeting transitive condition, is a fuzzy equivalence relation. With $R^*$ and a threshold value, linguistics data sequences are clustered into clusters. Therefore, after obtaining $R^*$, the appropriate threshold value $\lambda$ should be identified. Furthermore, $R^*$ can replace for $R^+$ in clustering.

Definition 7. Suppose that $R^-_A$ is a binary relation on $S$, $\lambda_A$ can be defined as $R^-_A = \{(x, y) | r(x, y) \geq \lambda, \forall x, y \in S\}$, where $0 \leq \lambda \leq 1$.

According to definition 7, $R^* = \{(i, j)|\mu_{R^*}(i, j) \geq \lambda\}$ and $0 \leq \lambda \leq 1$. By $R^-_A$, the clustering guideline will be presented as follow.

$$
\mu_{R^-}(i, j) = \mu_{R^+}(i, j) \geq \lambda, \text{then } i\text{th sequence and } j\text{th sequence belong to a same cluster; otherwise, the two sequences will be members of two separate clusters.}
$$

Because $R^*$ is a fuzzy equivalence relation, then $R^-_A$ is an equivalence relation. As $\lambda = 0$, $R_a$ classifies $m$ fuzzy data sequences into the same cluster. In contrast, when $\lambda = 1$, normally $R^-_A$ classifies $m$ fuzzy data sequences into $m$ separate clusters. A validation index (VI) is calculated to identify the appropriate value of $\lambda$.

Definition 8. A cluster of $S$ shows a group of separate subsets, like $\{S_1, S_2, \ldots, S_n\}$, in a way that the union of these subsets comprises the entire set $S$. In other words, $S_1 \cup S_2 \cup \ldots \cup S_n = S$ and $S_i \cup S_j = \emptyset, \forall i \neq j$. 5
Definition 9. Assume that \( P_{R_3} \) is a cluster of \( S \) resulted from an equivalence relation \( R_2 \). Define the elements \( x, y \in S \) to be in a subset \( S_1 \) of \( S \) as \((x, y) \in R_3 \), and \( S_1 \in P_{R_3} \).

Step 10. According to the validation index of partition \( P_{R_4} \) (for different \( \lambda \)), the best cluster is defined as

\[
VI = \max \sum_{i=1}^{m} \sum_{j=1}^{m} (\alpha(i,j) \times \mu_R^i(i,j) + (1 - \alpha(i,j)) \times (1 - \mu_R^i(i,j)))
\]

where

\[
\alpha(i,j) =
\begin{cases} 
1, & \text{if } i, j \in \hat{S} \in P_{R_4} \setminus \hat{S}
\end{cases}
\]

\[
0, & \text{if } i \in \hat{S}, j \in \hat{S}, \hat{S} \in P_{R_4} \setminus \hat{S} \neq \hat{S}.
\]

Partitioning data results in inter-cluster (between clusters) and intra-cluster (within cluster) relations. Normally, a specific clustering is preferred to the other one when it has a low inter-cluster relation and a high intra-cluster relation [33]. It is clear that when the number of data rises, the number of clusters increases which brings more inter-cluster and intra-cluster relations. Validation index considers these two relations and a higher validation index shows a better partitioning which means a higher intra-cluster relation between objects and a lower inter-cluster relation between objects. There is no upper bound for validation indices and they rise by the number of objects.

2.3. Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA) is a non-parametric fractional mathematical programming approach which divides the weighted sum of outputs by the weighted sum of inputs to measure the relative efficiency of a homogeneous group of decision-making units. Primarily, DEA is run to compare the relative efficiency of a group of homogeneous set of decision-making units and implemented in many contexts such as airports, hospitals, university departments, schools, industries and sectors, banks, products and services, and technologies [36; 8].

2.3.1. The CCR Model

Consider a group of \( n \) DMUs where each \( DMU_j, (j = 1, ..., n) \) generates \( s \) different outputs utilizing \( m \) different inputs that are denoted by \( y_{rj}, (r = 1, ..., s) \) and \( x_{ij}, (i = 1, ..., m) \) respectively. We assume \( x_{ij} \) and \( y_{rj} \) are all positive [37]. For every \( DMU_j \) under evaluation, the efficiency score \( E_j \) can be calculated by the CCR input oriented multiplier model (16):

\[
\text{Max } E_j = \sum_{r=1}^{s} u_{r} y_{r0}
\]

Subject to:

\[
\sum_{r=1}^{m} u_{r} y_{rj} - \sum_{i=1}^{s} v_{i} x_{ij} \leq 0, \quad j = 1, 2, ..., n.
\]

\[
\sum_{i=1}^{m} v_{i} x_{i0} = 1
\]

\[
u_{r} \geq 0, \quad v_{i} \geq 0, \quad r = 1, ..., s, \quad i = 1, ..., m.
\]

where

\[
\begin{align*}
\sum_{r=1}^{s} u_{r} y_{rj} & \leq 1, \\
v_{i} x_{ij} & \leq 1, \\
u_{r} & \geq 0, \quad v_{i} \geq 0, \quad r = 1, ..., s, \quad i = 1, ..., m.
\end{align*}
\]

The conventional DEA models such as CCR, need exact (crisp) values for inputs and outputs; however, crisp data might not always be attainable in real-world situations. Sometimes in real settings, inputs and outputs contain missing data, judgmental data, or predictive data or imprecise and vague data. One way to manage uncertain input and output data is to use fuzzy numbers to extend DEA models [38; 39; 40].

2.3.2. Common Set of Weights (CSW)

As CCR model calculates different set of weights for each DMU to maximize its relative efficiency, thus ranking efficient DMUs is not possible with this model. Scholars solved this problem using various methods [41] and it is noteworthy that one of the most popular techniques for assessing and ranking efficient DMUs is common set of weights (CSW) technique [42]. Different common weight models are developed including maximizing the efficiency score of the set of efficient DMUs using CWA-methodology [41], utilizing an auxiliary vector to transform the multi-objective fractional linear programming (MOFP) [43], regression analysis and nonlinear regression models [14], considering ideal and anti-ideal DMUs ([7;44]), goal programming (GP) [45; 46; 47], compromise programming [48], and displaced ideal methodology [49].
2.3.3. Ideal Point Method

Definition 10. The (virtual) ideal DMU is a DMU that its inputs are at the minimum level and its outputs are at the maximum level among all of the DMUs. The ideal DMU is shown with $\overline{DMU} = (\overline{x}, \overline{y})$ then $\overline{x}_i = \min \{x_{ij} | j = 1, ..., n\}$, $(i = 1, ..., m)$ and $\overline{y}_r = \max \{y_{rij} | j = 1, ..., n\}$, $(r = 1, ..., s)$. Here $\overline{x}$ denotes the inputs and $\overline{y}$ denotes the outputs of the ideal DMU. The prominent feature of the ideal point method compared to the others, is that it is always feasible [50]. Then, the CSW-IP model [44] is as (17).

$$\begin{align*}
\text{Min} & \quad \sum_{i=1}^{m} \sum_{j=1}^{n} v_i x_{ij} - \sum_{i=1}^{m} v_i \overline{x}_{\min} \\
\text{subject to:} & \quad \sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} v_r y_{rj} \geq 0, \quad j = 1, 2, \ldots, n \\
& \sum_{i=1}^{m} v_i \overline{x}_{\min} = 1, \quad \sum_{r=1}^{s} u_r \overline{y}_{\max} = 1, \\
& v_i u_r \geq \epsilon > 0, i = 1, \ldots, m, r = 1, \ldots, s.
\end{align*}$$

3. An Alternative Fuzzy CSW Considering the Ideal Point

From among different types of fuzzy numbers, triangular fuzzy numbers (TFNs) are more popular. In the sequel, the values of inputs and outputs of DMUs will be triangular fuzzy numbers.

In the simplest situation where there are one input and one output, the efficiency frontier of fuzzy CCR model will be as Figure 1. In Figure 1.a, the DMU consumes fuzzy input $\overline{x} = (x', x^m, x^u)$ to produce $y_o$. When the input $x_o$ is a crisp value and output of the DMU is $\overline{y} = (y', y^m, y^u)$, the efficiency frontier will be as Figure 1.b.

Figure 1. Efficiency frontier of fuzzy CCR

In Figure 2, the vertical axis shows the weighted sum of $s$ outputs and the horizontal axis depicts the weighted sum of $m$ inputs. Line “OO” is an ideal line (frontier) representing that all the DMUs on this frontier satisfy the constraint of the weighted sum of $s$ outputs equals to the weighted sum of $m$ inputs. For a given set of weights $u_r$, $(r = 1, \ldots, s)$ and $v_i$, $(i = 1, \ldots, m)$, the gap between points $M$ and ideal point is $\sum_{i=1}^{n} v_i x_{Mj} - \sum_{r=1}^{s} v_r y_{Mj}$ on the horizontal axis and $\sum_{r=1}^{s} u_r \overline{y}_{\max} - \sum_{i=1}^{n} u_i y_{ij}$ on the vertical axis. Likewise, for points $N$ and $L$, the gaps will be calculated.

Figure 2. Gap analysis for DMUs below the virtual ideal DMU

The objective is to calculate an optimal set of weights $u_r^*(r = 1, \ldots, s)$ and $v_i^*(i = 1, \ldots, m)$ so that all DMUs below the ideal frontier could be as close to their ideal point $(\overline{DMU})$ on the ideal frontier, as possible. In other words, implementing the optimal weights will minimize the total gaps between all DMUs and the ideal point.
Therefore, the common weight model using ideal
point method with fuzzy inputs and outputs (CSW-FIP) is proposed as (18):

\[
\begin{align*}
&\text{Min} \sum_{j=1}^{n} \left[ \sum_{i=1}^{m} v_i \bar{x}_{ij} - \sum_{i=1}^{m} v_i \bar{x}_{\min} \right] \\
&+ \sum_{j=1}^{n} \left[ \sum_{r=1}^{s} u_r \bar{y}_{\max} - \sum_{r=1}^{s} u_r \bar{y}_r \right] \\
\text{Subject to:} \\
&\sum_{i=1}^{m} v_i \bar{x}_{ij} - \sum_{r=1}^{s} u_r \bar{y}_r \geq 0, \quad j = 1, 2, ..., n \\
&\sum_{i=1}^{m} v_i \bar{x}_{\min} = 1, \quad \sum_{r=1}^{s} u_r \bar{y}_{\max} = 1, \\
&v_i, u_r \geq \varepsilon > 0, \quad i = 1, ..., m, \quad r = 1, ..., s.
\end{align*}
\]

(18)

where, ‘\(\geq\)’ indicates the fuzziness. Now let replace \(\bar{x}_{ij}\) with \(\sum_{i=1}^{m}(v_i \bar{x}_{ij}', v_i \bar{x}_{ij}^m, v_i \bar{x}_{ij}^u)\), \(\bar{y}_r\) with \(\sum_{i=1}^{n}(u_r \bar{y}_r^l, u_r \bar{y}_r^m, u_r \bar{y}_r^u)\), \(\bar{x}_{\min}\) with \(\sum_{i=1}^{m}(v_i \bar{x}_{\min}', v_i \bar{x}_{\min}^m, v_i \bar{x}_{\min}^u)\), and \(\bar{y}_{\max}\) with \(\sum_{r=1}^{s}(u_r \bar{y}_r^l, u_r \bar{y}_r^m, u_r \bar{y}_r^u)\).

**Min** \[
\begin{align*}
\sum_{j=1}^{n} \left[ \sum_{i=1}^{m} (v_i \bar{x}_{ij}', v_i \bar{x}_{ij}^m, v_i \bar{x}_{ij}^u) \\
- \sum_{i=1}^{m} (v_i \bar{x}_{\min}', v_i \bar{x}_{\min}^m, v_i \bar{x}_{\min}^u) \right] \\
+ \sum_{j=1}^{n} \left[ \sum_{r=1}^{s} (u_r \bar{y}_r^l, u_r \bar{y}_r^m, u_r \bar{y}_r^u) \right] \\
- \sum_{r=1}^{s} (u_r \bar{y}_r^l, u_r \bar{y}_r^m, u_r \bar{y}_r^u) \right] \\
\text{Subject to:} \\
\sum_{i=1}^{m} (v_i \bar{x}_{ij}', v_i \bar{x}_{ij}^m, v_i \bar{x}_{ij}^u) \\
- \sum_{i=1}^{m} (v_i \bar{x}_{\min}', v_i \bar{x}_{\min}^m, v_i \bar{x}_{\min}^u) \\
\geq (0,0,0), \quad j = 1, 2, ..., n \\
\sum_{i=1}^{m} (v_i \bar{x}_{\min}', v_i \bar{x}_{\min}^m, v_i \bar{x}_{\min}^u) = (1,1,1), \\
\sum_{r=1}^{s} (u_r \bar{y}_r^l, u_r \bar{y}_r^m, u_r \bar{y}_r^u) = (1,1,1), \\
v_i, u_r \geq \varepsilon > 0, \quad i = 1, ..., m, \quad r = 1, ..., s.
\end{align*}
\]

(19)

Model (19) is a fuzzy multiplier linear programming and cannot be solved by standard linear programming solvers because of the fuzzy numbers. Four categories of approaches to solve fuzzy DEA are (1) the tolerance technique, (2) the \(\alpha\)-cut technique, (3) the fuzzy ranking technique, and (4) the possibility technique. The \(\alpha\)-cut technique is more commonly used to solve fuzzy DEA model [40; 51]. While more linear programs should be solved choosing the \(\alpha\)-cut approach, we selected this technique because of two main reasons: (1) this approach allows modelling crisp values of inputs and outputs as triangular fuzzy numbers if any of those data are found to be crisp, and (2) owing to the extension principle, this approach transforms the fuzzy DEA model into a crisp DEA model which is easily solved as a linear program [52].

Thus, this method is applied to solve (19). First, introduce \(\alpha\)-cuts of the objective function and constraint in (19). Then Model (19) transforms into (20):

\[
\begin{align*}
\text{Min} \sum_{i=1}^{m} \left[ \sum_{j=1}^{n} v_i (ax_{ij}^m + (1 - \alpha) ax_{ij}^l) \right] \\
- \sum_{i=1}^{m} v_i (ax_{ij}^u + (1 - \alpha) ax_{ij}^l) \\
\text{Subject to:} \\
\sum_{j=1}^{n} v_i (ax_{ij}^m + (1 - \alpha) ax_{ij}^l) \\
- \sum_{i=1}^{m} u_r (ay_{rj}^m + (1 - \alpha) ay_{rj}^l) \\
\geq (0,0,0), \quad j = 1, 2, ..., n \\
\sum_{i=1}^{m} (v_i \bar{x}_{\min}', v_i \bar{x}_{\min}^m, v_i \bar{x}_{\min}^u) = (1,1,1), \\
\sum_{r=1}^{s} (u_r \bar{y}_r^l, u_r \bar{y}_r^m, u_r \bar{y}_r^u) = (1,1,1), \\
v_i, u_r \geq \varepsilon > 0, \quad i = 1, ..., m, \quad r = 1, ..., s.
\end{align*}
\]

(20)

8
\[
\left( \sum_{r=1}^{s} u_r(\alpha \overline{y}_{\text{max}}(r) + (1 - \alpha)\overline{y}^u_{\text{max}}(r), \alpha \overline{y}_{\text{max}}(r) + (1 - \alpha)\overline{y}^n_{\text{max}}(r)) \right) = [1, 1],
\]

\[ v_i, u_r \geq \varepsilon > 0, \quad i = 1, \ldots, m, \quad r = 1, \ldots, s, \quad j = 1, 2, \ldots, n \]

All the coefficients in (20) are intervals which imply that Model (20) is a distant problem. The following steps are taken to solve (20):

Step 1. Consider the variables \( \hat{x}_{ij}, \hat{x}_{ij}^\text{min}, \hat{y}_{rj} \) and \( \hat{y}_{rj}^\text{max} \) such that:

\[ \hat{x}_{ij} \in [(\alpha x_{ij}^m + (1 - \alpha)x_{ij}^l, \alpha x_{ij}^m + (1 - \alpha)x_{ij}^l), \forall i, j], \]

\[ \hat{x}_{ij}^\text{min} \in [(\alpha \overline{x}_{\text{min}}^m + (1 - \alpha)\overline{x}_{\text{min}}^l, \alpha \overline{x}_{\text{min}}^m + (1 - \alpha)\overline{x}_{\text{min}}^l), \forall i, j], \]

\[ \hat{y}_{rj} \in [(\alpha y_{rj}^m + (1 - \alpha)y_{rj}^l, \alpha y_{rj}^m + (1 - \alpha)y_{rj}^l), \forall r, j], \]

\[ \hat{y}_{rj}^\text{max} \in [(\alpha \overline{y}_{\text{max}}^m + (1 - \alpha)\overline{y}_{\text{max}}^l, \alpha \overline{y}_{\text{max}}^m + (1 - \alpha)\overline{y}_{\text{max}}^l), \forall r, j]. \]

Step 2. Substituting these variables in (20) then:

\[
\text{Min} \sum_{j=1}^{n} \sum_{i=1}^{m} v_i \hat{x}_{ij} - \sum_{i=1}^{m} \sum_{r=1}^{s} v_i \hat{x}_{ij}^\text{min}
\]

\[
+ \sum_{j=1}^{n} \sum_{r=1}^{s} u_r \hat{y}_{rj}^\text{max} - \sum_{r=1}^{s} u_r \hat{y}_{rj} \]

Subject to:

\[
\sum_{i=1}^{m} v_i \hat{x}_{ij} - \sum_{r=1}^{s} u_r \hat{y}_{rj} \geq 0, \quad j = 1, 2, \ldots, n
\]

\[
\sum_{i=1}^{m} v_i \hat{x}_{ij}^\text{min} = 1, \quad \sum_{r=1}^{s} u_r \hat{y}_{rj}^\text{max} = 1,
\]

\[
\alpha x_{ij}^m + (1 - \alpha)x_{ij}^l \leq \hat{x}_{ij} \leq \alpha x_{ij}^m + (1 - \alpha)x_{ij}^l, \forall i, j,
\]

\[
\alpha \overline{x}_{\text{min}}^m + (1 - \alpha)\overline{x}_{\text{min}}^l \leq \hat{x}_{ij}^\text{min} \leq \alpha \overline{x}_{\text{min}}^m + (1 - \alpha)\overline{x}_{\text{min}}^l, \forall i, j,
\]

\[
\alpha y_{rj}^m + (1 - \alpha)y_{rj}^l \leq \hat{y}_{rj} \leq \alpha y_{rj}^m + (1 - \alpha)y_{rj}^l, \forall r, j,
\]

\[
\alpha \overline{y}_{\text{max}}^m + (1 - \alpha)\overline{y}_{\text{max}}^l \leq \hat{y}_{rj}^\text{max} \leq \alpha \overline{y}_{\text{max}}^m + (1 - \alpha)\overline{y}_{\text{max}}^l, \forall r, j.
\]

Model (21) is a crisp nonlinear programming problem where \( \alpha \in (0, 1] \) is a parameter and the optimal solution for different values of \( \alpha \) can be found.

Definition 11. A DMU\(_j\) is efficient at given \( \alpha \in (0, 1] \) if \( E_j^{\alpha*} = 1 \).

Nonlinear programming (21) is converted into linear programming problem (LP).

Step 3. Substitute the variables \( \hat{x}_{ij} = v_i \hat{x}_{ij}, \hat{y}_{rj} = u_r \hat{y}_{rj}, \hat{x}_{ij}^\text{min} = v_i \hat{x}_{ij}^\text{min}, \hat{y}_{rj}^\text{max} = u_r \hat{y}_{rj}^\text{max} \) in (21) to reduces to an LP as (22).

\[
\text{Min} \sum_{j=1}^{n} \sum_{i=1}^{m} v_i (\alpha x_{ij}^m + (1 - \alpha)x_{ij}^l), \forall i, j,
\]

\[
\sum_{r=1}^{s} u_r (\alpha \overline{x}_{\text{min}}^m + (1 - \alpha)\overline{x}_{\text{min}}^l), \forall r, j,
\]

\[
u_i, u_r \geq \varepsilon > 0, i = 1, \ldots, m, r = 1, \ldots, s. \]

Model (22) is a crisp linear programming problem where \( \alpha \in (0, 1] \) is a parameter and we find an optimal solution for each value of \( \alpha \). The variable transformations in step (3), \( \hat{x}_{ij}, \hat{y}_{rj} \) are the weighted inputs and the weighted outputs of DMU\(_j\), in which \( v_i, u_r \) are the common weights generated by the proposed ideal point method. Therefore, the efficiency score of DMU\(_j\) with CSW is:

\[
E_j = \frac{\sum_{i=1}^{m} \hat{y}_{rj}/\sum_{i=1}^{m} \hat{x}_{ij}}{\sum_{i=1}^{m} \hat{x}_{ij}} \quad (23)
\]

Since the available real-world data of inputs and outputs are uncertain in most cases, therefore, the proposed common set of weights approach is more appropriate in such situations. Parameter \( \alpha \) can adopt different values between \([0, 1]\). When the value of \( \alpha \) is smaller, the support of the membership function becomes wider. The wider the support of the membership function, the higher the uncertainty. That is, \( \alpha = 1 \) shows the most certain efficiency measurement while \( \alpha = 0 \) shows the most uncertain efficiency.
measurement. On the other hand, efficiency values obtained at $\alpha = 1$ would equal the efficiency score of DMUs when their inputs and outputs are crisp values. From risk management perspective, risk averse decision makers prefer higher values of $\alpha$ to have the higher certainty while risk takers prefer lower values of $\alpha$.

4. An Application in Banking

4.1. Data and Variables Illustration

To follow proposed method for clustering and evaluating efficiency using DEA, several environmental criteria have been considered. The identification and classification of these criteria have been conducted based on both previous research such as Yang [53], Shyu and Chiang [54], Paradi et al. [55], Paradi and Zhu [56], Frederick [57], Katircioglu et al. [58], Al Shaheer et al. [59] and experts’ opinions. Effective environmental criteria are shown in Table 2.

Selecting inputs and outputs is a critical part of performance analysis, as it includes evaluating different aspects of branch performance. Certainly, since banks deliver several products and services, it is too difficult to consider the whole range of banking activities. The European banking industry considers the intermediation approach which acknowledges intermediation as the core activity. It means that banks do not produce loan and deposit services, instead, deposits are dealt with as inputs while, loans and investments are considered as outputs. Therefore, to get a clear understanding of the proposed methodology, this paper evaluates efficiency of 100 DMUs (Melli bank branches in the province of Isfahan) with three fuzzy inputs and two fuzzy outputs. The data of fuzzy inputs and outputs are positive TFNs as shown in Table 3.

<table>
<thead>
<tr>
<th>Input variable</th>
<th>Output variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$ Total costs</td>
<td>$y_1$ Total credits (facilities) provided to customers</td>
</tr>
<tr>
<td>$x_2$ Total deposits</td>
<td>$y_2$ Total income</td>
</tr>
<tr>
<td>$x_3$ Total assets</td>
<td>-</td>
</tr>
</tbody>
</table>

4.2. Results and Discussion

Clustering algorithm has been formulated in MS Excel 2016, and all DEA computation were done by MATLAB version R2016b. A group of five experts who are the senior managers in the Melli bank of province of Isfahan participated in this study. They evaluated 100 branches of Melli bank against the 31 criteria to cluster them. Table 4 shows the score of alternatives against the criteria which are attributed by the decision maker No.1. For example, alternative A1 has the score of “Very High” in C1. Table 5 shows the equivalent triangular fuzzy numbers of scores in Table 4. For example, the score of “Very High” for A1 against C1 in Table 4, is equivalent to $(0.8,1,1)$ in Table 5.

Table 4. The linguistic assessments of alternatives against criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
<th>C9</th>
<th>C10</th>
<th>C11</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>VH</td>
<td>BMH</td>
<td>VH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>H</td>
<td>VL</td>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>VH</td>
<td>VH</td>
<td>VH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A100</td>
<td>H</td>
<td>BMH</td>
<td>BMH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Assessments of alternatives against criteria with TFNs

<table>
<thead>
<tr>
<th>Criteria</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
<th>C9</th>
<th>C10</th>
<th>C11</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>0.8</td>
<td>0.5</td>
<td>0.6</td>
<td>0.8</td>
<td>0.8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.6</td>
<td>0.8</td>
<td>1</td>
</tr>
<tr>
<td>A2</td>
<td>0.6</td>
<td>0.6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0.6</td>
<td>0.8</td>
<td>1</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>A3</td>
<td>0.8</td>
<td>0.8</td>
<td>1</td>
<td>0.8</td>
<td>1</td>
<td>0.8</td>
<td>1</td>
<td>1</td>
<td>0.6</td>
<td>0.8</td>
<td>1</td>
</tr>
<tr>
<td>A100</td>
<td>0.6</td>
<td>0.6</td>
<td>0.5</td>
<td>0.6</td>
<td>0.8</td>
<td>0.5</td>
<td>0.6</td>
<td>0.8</td>
<td>1</td>
<td>0.5</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Table 6 reports details about the maximum distance between upper and lower bounds of $q^*$ fuzzy number of all sequences.

Table 6. Distance between upper and lower boundaries of $q^*$ fuzzy number of all sequences

<table>
<thead>
<tr>
<th>$d(T_a, T_b)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
</tr>
<tr>
<td>0.7</td>
</tr>
<tr>
<td>C12</td>
</tr>
<tr>
<td>0.7</td>
</tr>
<tr>
<td>C23</td>
</tr>
<tr>
<td>V 1.8</td>
</tr>
</tbody>
</table>

* Note: C: Criteria; V: Value

Table 7 shows the initial relation matrix of all sequences and Table 8 shows the fuzzy equivalence relations of the 100 branches’ preferences against all criteria that are obtained by the max–min transitive closure.
Table 2. Effective environmental criteria on bank branch efficiency

<table>
<thead>
<tr>
<th>Sub-criteria</th>
<th>Type of location</th>
<th>Rates of population growth in place</th>
<th>Size of population in place</th>
<th>Population density in place</th>
<th>Local unemployment rate</th>
<th>Production, distribution and business condition in place</th>
<th>Educational level in place</th>
<th>Branch age</th>
<th>Branch size</th>
<th>Physical work environment for employees</th>
<th>Competence and friendliness of bank personnel</th>
<th>Availability and convenient location of ATM services</th>
<th>Average employment income in the branch area</th>
<th>Number of branches in area according to population and business density</th>
<th>Available public transport nearby</th>
<th>Advertisement of branch</th>
<th>The diversity of rival banks branches in the area</th>
<th>Proximity to commercial venues and industrial settlements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Village</td>
<td>Rural district</td>
<td>City</td>
<td>District</td>
<td>County</td>
<td>Province</td>
<td>Metropolitan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VL</td>
<td>LO</td>
<td>BLM</td>
<td>M</td>
<td>BMH</td>
<td>H</td>
<td>VH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>* Note: BTW: Between low and medium; BMH: Between medium and high; BAH: Between average and high; BAG: Between small and average; BLM: Between low and medium; BFA: Between few and average; M: Medium; A: Average; VL: Very low; VH: Very high; VB: Very bad; S: Small; VS: Very small; L: Large; VG: Very good; LV: Very large; B: Bad; G: Good; LO: Low; MA: Many; SM: So many; F: Few; H: Very few; P: Public; N-P: None-public; E: Employee; B: Business; I: Industrial; G: Government; IN: International;</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 9. The partitioning of the 100 branches on different intervals of $\lambda$s

<table>
<thead>
<tr>
<th>$\lambda$ Level</th>
<th>Clusters</th>
<th>Clustering arrangements</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.875,1]</td>
<td>10</td>
<td><img src="image" alt="Clusters" /></td>
</tr>
<tr>
<td>(0.806,0.875]</td>
<td>8</td>
<td><img src="image" alt="Clusters" /></td>
</tr>
<tr>
<td>(0.623,0.806]</td>
<td>7</td>
<td><img src="image" alt="Clusters" /></td>
</tr>
<tr>
<td>(0.511,0.623]</td>
<td>6</td>
<td><img src="image" alt="Clusters" /></td>
</tr>
<tr>
<td>(0.496,0.511]</td>
<td>5</td>
<td><img src="image" alt="Clusters" /></td>
</tr>
<tr>
<td>(0.453,0.496]</td>
<td>4</td>
<td><img src="image" alt="Clusters" /></td>
</tr>
<tr>
<td>(0.446,0.453]</td>
<td>3</td>
<td><img src="image" alt="Clusters" /></td>
</tr>
<tr>
<td>(0.409,0.446]</td>
<td></td>
<td><img src="image" alt="Clusters" /></td>
</tr>
</tbody>
</table>
Table 10. Input and output variables (10 Rials)

<table>
<thead>
<tr>
<th>No. of DMU</th>
<th>Total costs ($x_1$)</th>
<th>Total deposits ($x_2$)</th>
<th>Total assets ($x_3$)</th>
<th>Total credits (facilities provided to customers) ($y_1$)</th>
<th>Total income ($y_2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(4422061610, 44238086139, 44293165547)</td>
<td>(5336754127, 53384224106, 53439303514)</td>
<td>(4090645580, 4108115559, 416194967)</td>
<td>(33074131661, 33091853640, 33146663048)</td>
<td>(5162415170, 5179885149, 5234964557)</td>
</tr>
<tr>
<td>2</td>
<td>(306389146922, 30706879603, 30761956311)</td>
<td>(7715565166, 7717325145, 7722840553)</td>
<td>(8401167482, 8416857461, 8473716869)</td>
<td>(2541839856, 2543600935, 25491089243)</td>
<td>(9046165131, 9063653110, 9118714518)</td>
</tr>
<tr>
<td>3</td>
<td>(284467092490, 28479328216, 28528732328)</td>
<td>(48672873742, 48705049018, 48755043130)</td>
<td>(9858417419, 9891042695, 9940586807)</td>
<td>(2829881993, 28331445259, 28380983971)</td>
<td>(742489621, 7429385571, 742986921)</td>
</tr>
<tr>
<td>100</td>
<td>(22743115413, 22833439024, 2292634817)</td>
<td>(49413980211, 49509223722, 49608472615)</td>
<td>(7592286774, 7682622885, 7781806178)</td>
<td>(38209628060, 3829943571, 38399147464)</td>
<td>(5875610531, 5965929042, 6065129935)</td>
</tr>
</tbody>
</table>

*Note: VI: Validation index

Table 11. Efficiency values for 100 bank branches

<table>
<thead>
<tr>
<th>Cluster</th>
<th>DMUs</th>
<th>Efficiency</th>
<th>Rank</th>
<th>Efficiency</th>
<th>Rank</th>
<th>Efficiency</th>
<th>Rank</th>
<th>Efficiency</th>
<th>Rank</th>
<th>Efficiency</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>0.001354</td>
<td>92</td>
<td>0.207159</td>
<td>74</td>
<td>0.143923</td>
<td>96</td>
<td>0.17443</td>
<td>89</td>
<td>0.145748</td>
<td>97</td>
</tr>
<tr>
<td>A3</td>
<td>1</td>
<td>0.001346</td>
<td>93</td>
<td>0.226707</td>
<td>65</td>
<td>0.19307</td>
<td>98</td>
<td>0.191557</td>
<td>87</td>
<td>0.139802</td>
<td>100</td>
</tr>
<tr>
<td>A11</td>
<td>1</td>
<td>0.007334</td>
<td>75</td>
<td>0.631695</td>
<td>2</td>
<td>0.840064</td>
<td>1</td>
<td>0.681175</td>
<td>2</td>
<td>0.731043</td>
<td>1</td>
</tr>
<tr>
<td>A12</td>
<td>1</td>
<td>0.001139</td>
<td>94</td>
<td>0.143923</td>
<td>96</td>
<td>0.17443</td>
<td>89</td>
<td>0.145748</td>
<td>97</td>
<td>0.159629</td>
<td>77</td>
</tr>
<tr>
<td>A15</td>
<td>1</td>
<td>0.001355</td>
<td>95</td>
<td>0.207733</td>
<td>72</td>
<td>0.338402</td>
<td>29</td>
<td>0.211637</td>
<td>74</td>
<td>0.229835</td>
<td>52</td>
</tr>
<tr>
<td>A16</td>
<td>1</td>
<td>0.000977</td>
<td>98</td>
<td>0.19037</td>
<td>98</td>
<td>0.191557</td>
<td>87</td>
<td>0.139802</td>
<td>100</td>
<td>0.153487</td>
<td>79</td>
</tr>
<tr>
<td>A19</td>
<td>1</td>
<td>0.00142</td>
<td>89</td>
<td>0.193854</td>
<td>81</td>
<td>0.262406</td>
<td>53</td>
<td>0.193904</td>
<td>81</td>
<td>0.210849</td>
<td>63</td>
</tr>
<tr>
<td>A20</td>
<td>1</td>
<td>0.000779</td>
<td>100</td>
<td>0.159181</td>
<td>90</td>
<td>0.200615</td>
<td>85</td>
<td>0.160176</td>
<td>90</td>
<td>0.175558</td>
<td>73</td>
</tr>
<tr>
<td>A23</td>
<td>1</td>
<td>0.0034</td>
<td>78</td>
<td>0.48387</td>
<td>8</td>
<td>0.641223</td>
<td>3</td>
<td>0.506921</td>
<td>9</td>
<td>0.549988</td>
<td>54</td>
</tr>
<tr>
<td>A29</td>
<td>1</td>
<td>0.00184</td>
<td>85</td>
<td>0.257912</td>
<td>46</td>
<td>0.350353</td>
<td>25</td>
<td>0.256922</td>
<td>49</td>
<td>0.278582</td>
<td>33</td>
</tr>
<tr>
<td>A30</td>
<td>1</td>
<td>0.000942</td>
<td>99</td>
<td>0.189651</td>
<td>82</td>
<td>0.236583</td>
<td>68</td>
<td>0.189776</td>
<td>83</td>
<td>0.207692</td>
<td>64</td>
</tr>
<tr>
<td>A32</td>
<td>1</td>
<td>0.001618</td>
<td>87</td>
<td>0.242087</td>
<td>54</td>
<td>0.3009</td>
<td>36</td>
<td>0.241098</td>
<td>57</td>
<td>0.257884</td>
<td>34</td>
</tr>
<tr>
<td>A47</td>
<td>1</td>
<td>0.002073</td>
<td>83</td>
<td>0.299536</td>
<td>32</td>
<td>0.404236</td>
<td>16</td>
<td>0.30207</td>
<td>35</td>
<td>0.326835</td>
<td>24</td>
</tr>
<tr>
<td>A51</td>
<td>1</td>
<td>0.00187</td>
<td>84</td>
<td>0.241126</td>
<td>56</td>
<td>0.306979</td>
<td>33</td>
<td>0.240711</td>
<td>53</td>
<td>0.267713</td>
<td>37</td>
</tr>
<tr>
<td>A60</td>
<td>1</td>
<td>0.001054</td>
<td>95</td>
<td>0.288125</td>
<td>71</td>
<td>0.272552</td>
<td>49</td>
<td>0.296113</td>
<td>76</td>
<td>0.223777</td>
<td>57</td>
</tr>
<tr>
<td>A62</td>
<td>1</td>
<td>0.005274</td>
<td>76</td>
<td>0.476541</td>
<td>9</td>
<td>0.59631</td>
<td>50</td>
<td>0.506925</td>
<td>8</td>
<td>0.557729</td>
<td>4</td>
</tr>
<tr>
<td>A72</td>
<td>1</td>
<td>0.002684</td>
<td>80</td>
<td>0.343844</td>
<td>25</td>
<td>0.388382</td>
<td>20</td>
<td>0.351731</td>
<td>22</td>
<td>0.385361</td>
<td>15</td>
</tr>
<tr>
<td>A75</td>
<td>1</td>
<td>0.00137</td>
<td>90</td>
<td>0.181209</td>
<td>84</td>
<td>0.245692</td>
<td>62</td>
<td>0.181462</td>
<td>84</td>
<td>0.199797</td>
<td>67</td>
</tr>
<tr>
<td>A77</td>
<td>1</td>
<td>0.003417</td>
<td>77</td>
<td>0.406741</td>
<td>14</td>
<td>0.571091</td>
<td>7</td>
<td>0.433106</td>
<td>15</td>
<td>0.471703</td>
<td>9</td>
</tr>
<tr>
<td>A78</td>
<td>1</td>
<td>0.001481</td>
<td>88</td>
<td>0.211616</td>
<td>70</td>
<td>0.26892</td>
<td>51</td>
<td>0.21477</td>
<td>71</td>
<td>0.235858</td>
<td>50</td>
</tr>
<tr>
<td>A83</td>
<td>1</td>
<td>0.001026</td>
<td>96</td>
<td>0.156337</td>
<td>92</td>
<td>0.204442</td>
<td>83</td>
<td>0.15817</td>
<td>92</td>
<td>0.172957</td>
<td>74</td>
</tr>
<tr>
<td>A85</td>
<td>1</td>
<td>0.002368</td>
<td>82</td>
<td>0.288876</td>
<td>35</td>
<td>0.363035</td>
<td>22</td>
<td>0.29717</td>
<td>37</td>
<td>0.334291</td>
<td>22</td>
</tr>
</tbody>
</table>

Efficiency score of DMUs using the proposed common weight model (22) at different levels of α is shown in Table 11.
The efficiency of branches in each cluster are measured with the developed CSW-F-IP model and all 100 branches are ranked at different levels of α. Results show that a branch with high efficiency may produce lower outputs than the branch with lower efficiency when receiving the same inputs or even in some cases more inputs. However, it is ranked higher because of the environmental factors that have made heterogeneous conditions for branches in different clusters.

5. Advantages and Limitations

After introducing the useful clustering methods, the fuzzy equivalence relation method is used due to the following reasons: the existence of uncertain data, the large number of DMUs, the lack of clarity in the number of clusters at the beginning of the work, and solving the problem of overlapping clusters. This method eventually leads to the clustering of decision-making units in homogeneous groups by obtaining more appropriate and accurate answers. Each cluster includes those DMUs that have homogenous operational and environmental conditions which is the fundamental assumption to apply DEA. The proposed approach pays the highest attention to homogenizing the decision-making units and so provides fairer results for the managers and policy makers to make relevant decisions. This research contributes to the theory of DEA by 1) overcoming the heterogeneity of DMUs in terms of operational and environmental conditions through clustering with fuzzy equivalence relation, 2) considering fuzzy data for inputs and outputs which makes the DEA model nonlinear, 3) obtaining more information about the efficiency of DMUs, and 4) developing common set of weights for complete ranking of DMUs. The ideal point method has some merits over other approaches of finding CSW such as feasibility of the proposed model, possibility of defining a DMU as the best decision-making unit, and also its simplicity in contrast to the other methods. This study faces two limitations. First limitation relates to the number of inputs and outputs. Although there are quite a lot of DMUs under consideration, but the number of DMUs in some clusters are very few while for better discrimination among DMUs we need at least three times of the sum of inputs and outputs. To handle this limitation, we selected the most common factors as the proxies of input and output variables. Second limitation relates to the number of participant experts and their thorough knowledge on the operating environment of bank branches. Although there are many senior managers who are professional in banking industry, not all of them have extensive knowledge of environmental conditions of the branches. To overcome this barrier, we found five experts with snowball sampling which is enough for multi criteria decision-making situations.

6. Conclusion and Directions for Future Research

Conventional models of DEA accept that management of each DMU controls the inputs and outputs at his/her preference. Those models assume that all the DMUs are functioning under the same environmental situations. On the other hand, homogeneity of DMUs is a prerequisite for them. In reality, however, there might exist environmental inputs or outputs which DMU’s management is unable to control them. In many cases, it is essential to consider the environmental factors for performance evaluation. Those factors mainly relate to the operational environment of DMUs, which challenge the homogeneity presumption in DEA. In this study, we investigated the environmental conditions of bank branches to cluster them into homogenous decision-making units for a fairer efficiency analysis. To incorporate the uncertainty of human judgment in efficiency analysis, fuzzy
equivalence relation approach is used to cluster DMUs in terms of environmental and operational conditions. On the other hand, to overcome the weight flexibility of conventional DEA models and reach a complete ranking of all DMUs, we developed a fuzzy common set of weights model using ideal point method (CSWF-IP) to measure the efficiency of DMUs in each cluster. So, this paper contributes to knowledge of performance measurement by (a) clustering DMUs into homogeneous groups in terms of their operational and environmental conditions, and (b) improving discrimination among DMUs using the proposed common weight analysis with fuzzy data. Future studies can use multi-criteria decision-making methods to select the most important environmental factors for clustering. Finally, we recommend researchers to develop common weight models with goal programming technique and fuzzy data for non-discretionary factors.

Acknowledgement
The authors would like to thank the editor and two reviewers for their insightful comments, as a result, this paper has been improved substantially.

References


